VISVESVARAYA TECHNOLOGICAL UNIVERSITY
Belagavi-590018, Karnataka

A PROJECT REPORT
ON
PRODUCT REVIEW ANALYSIS

Submitted in partial fulfillment of requirements for the award of degree,

BACHELOR OF ENGINEERING

IN

COMPUTER SCIENCE & ENGINEERING

Submitted By:

1. SHASHANK G KAKADE	1MJ15CS133
2. SHASHIDHARA M	1MJ15CS135
3. SHASHIKUMAR N	1MJ15CS136
4. SHIVAKUMAR BHAT	1MJ15CS138

Under the Guidance of
Mrs. VANDANA
Assistant Professor, Department of Computer Science &Engineering

MVJ COLLEGE OF ENGINEERING Since 1982

(Affiliated to Visvesvaraya Technological University, Belagavi
Approved By AICTE, New Delhi,
Recognized by UGC under 2(f) & 12(B)
Accredited by NBA and NAAC)

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
MVJ COLLEGE OF ENGINEERING
BANGALORE-67

MVJ COLLEGE OF ENGINEERING

Whitefield, Near ITPB, Bangalore-67

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

COLLEGE OF ENGINEERING Since 1982

(Affiliated to Visvesvaraya Technological University, Belagavi
Approved By AICTE, New Delhi,
Recognized by UGC under 2(f) & 12(B)
Accredited by NBA and NAAC)

CERTIFICATE

This is to certify that the project work, entitled "**PRODUCT REVIEW ANALYSIS"** is a bona fide work carried out by

1. SHASHANK G KAKADE 1MJ15CS133
2. SHASHIDHARA M 1MJ15CS135
3. SHASHIKUMAR N 1MJ15CS136
4. SHIVAKUMAR BHAT 1MJ15CS138

in partial fulfillment for the award of degree of Bachelor of Engineering in Computer Science & Engineering of the Visvesvaraya Technological University, Belagavi during the academic year 2018-19. It is certified that all the corrections/suggestions indicated for Internal Assessment have been incorporated in the Report. The project report has been approved as it satisfies the academic requirements.

Signature of the Internal Guide	**Signature of the HOD**	**Signature of the Principal**

Signature of Internal Examiner	**Signature of External Examiner**

MVJ COLLEGE OF ENGINEERING

Whitefield, Near ITPB, Bangalore-67

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

**(Affiliated to Visvesvaraya Technological University, Belagavi
Approved By AICTE, New Delhi,
Recognized by UGC under 2(f) & 12(B)
Accredited by NBA and NAAC)**

DECLARATION

We, **Shashank G Kakade, Shashidhara M, Shashikumar N, Shivakumar Bhat** hereby declare that the entire work titled "**PRODUCT REVIEW ANALYSIS**" embodied in this project report has been carried out by us during the 8th semester of BE degree at MVJCE, Bangalore under the esteemed guidance of Mrs. Vandana, (Asst. Professor, Dept. of CSE, MVJCE) affiliated to Visvesvaraya Technological University, Belagavi. The work embodied in this dissertation work is original and it has not been submitted in part of full for any other degree in any University.

1. SHASHANK G KAKADE	1MJ15CS133	_____
2. SHASHIDHARA M	1MJ15CS135	_____
3. SHASHIKUMAR N	1MJ15CS136	_____
4. SHIVAKUMAR BHAT	1MJ15CS138	_____

Place:

Date:

PUBLICATION DETAILS

1. **Paper Name: -** Product Review Analysis using NLP (Natural Language Processing)
2. **Publication: -** ICAIIS at MVJ College of Engineering (International Conference on Innovations in Computing, Automation and Intelligent Information Systems)

ABSTRACT

Internet shopping is becoming more and more popular in modern era. The increase in its popularity has led to increase in customer reviews and opinion that a product receives. A customer who has to choose the right product among the huge varieties of products, depends heavily on the product reviews to make a purchase decision. With great volume of product reviews, it become difficult for customers to wade through all reviews to make an informed product choice. Nowadays customers look for features that can serve them specifically. But from the thousands of reviews, it is practically impossible for customers to identify the reviews which speak about the specific product feature. Hence there is a need for a system which analyses the review and gives the summarized data.

ACKNOWLEDGEMENT

The satisfaction and euphoria that accompany a successful completion of any task would be incomplete without the mention of people who made it possible, success is the epitome of hard work and perseverance, but steadfast of all is encouraging guidance.

So with gratitude we acknowledge all those whose guidance and encouragement served as beacon of light and crowned our effort with success.

We are thankful to our Principal **Dr. Nagaraj Sitaram,** for his encouragement and support throughout the project work.

We are also thankful to our beloved **HOD, Prof.I.Manimozi** for her incessant encouragement & all the help during the project work.

We consider it a privilege and honour to express our sincere gratitude to our guide **Mrs. Vandana, AP, Dept. of CSE** for her valuable guidance throughout the tenure of this project work, and whose support and encouragement made this work possible.

It's also a great pleasure to express our deepest gratitude to all the other faculty members of our department for their cooperation and constructive criticism offered, which helped us a lot during our project work.

Finally, we would like to thank all our family members and friends whose encouragement and support was invaluable.

Thanking you

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Project Overview

Popularity of online shopping has increased rapidly over the years. This growth is the result of enormous advantages, online shopping has over the conventional offline shopping practice. Some of the noted merits are, the time it saves to buy a list of products with few mouse clicks, freedom to determine affordable item by a straightforward comparison of prices, 24/7 availability of online stores and the comfort of the office or home. But it is difficult to get help from the professional sales staff to buy a product. To deal with this, merchant started providing sale related data for the products sold online. For customers, it was difficult to make decision about the product, with the details the product speaks about itself. As a solution to this, online merchants and e-commerce sites enabled forums, which allow customers to express their opinions and to get reviews.

Consumer reviews are significantly more reliable than the descriptions that come from manufacturers. This is because, a consumer review serves to explain what the product is about and how it works and these are given only after the customer's experience with the product. Even the badly written reviews will still give the reader some information on the products features and functions. Those that are better written will even show the potential buyer the benefits of using this product over another or perhaps pointing a negative feature. Consulting reviews is now a logical step in the purchasing cycle for all types of products and services and a customer who has to choose the right product among the huge varieties of products, depends heavily on the product reviews to make a purchase decision.

With great volume of product reviews, it become difficult for customers to wade through all reviews to make an informed product choice. Nowadays customers look for features that can serve them specifically. But from the thousands of reviews, it is practically impossible for customers to identify the reviews which speak about the specific product feature. Hence there is a need for a system which analyses the review and gives the summarized data.

1.2 Purpose of Project

Main purpose of the project is to analyse the reviews based on product features. The key module of this system is the product feature extraction module, which extracts product features from unstructured reviews. We propose a new model, which maps product features with the opinions. Scores will be assigned to each feature.

1.3 Scope of Project

The primary scope of project is to summarise the product reviews based on the comments given by the customer. The comments show the opinion of the user towards the product. These comments may be positive or negative. These comments may be in the form of sentences.

1.4 Problem Statement

Given a product review containing multiple features and varied opinions, the objective is to extract expressions of opinion describing a target feature and classify it as positive or negative.

1.5 Existing System

Pang and Lee suggested to remove objective sentences by extracting subjective ones. They proposed a text-categorization technique that is able to identify subjective content using minimum cut.

Xing Fang and Justin Zhan describes that the subjective contents are extracted, it consists of sentiment sentences which contain at least one positive or negative word.

These sentences are tokenized into separated English words. Depending on parts of speech in the words, corresponding tags are used.

In another work by Hu and Lui, they summarized customer reviews by polarities.

1.5.1 Disadvantages of Existing System

- The problem of imbalanced data distribution.
- Lacked Intuitive functionality.
- Most of the existing systems do not analyse the credibility of the reviews, many of them do not take into account the fake reviews as well of neutral reviews and in part apply the incorrect method of classification of the product.
- Complex User Interface which lacked convenience.

1.6 Proposed System

The system analyses the reviews based on product features. The key module of this system is the product feature extraction module, which extracts product features from unstructured reviews. Unstructured reviews mainly will be in the form of comments which show the opinion of the user towards the product. These comments may be positive or negative. These comments may be in the form of sentences. We propose a new model, which maps product features with the opinions. Scores will be assigned to each feature.

In order to gain the sentiment of the user, these sentences have to be segregated into words in which the adjectives, nouns and adverbs are processed using Natural Language Processing (NLP). Nouns will be considered as the opinion target or the feature and adjectives will be considered as opinion words. Based on the extraction of words, respective noun and adjective will be paired. So, opinion target and opinion word will be mapped. Finally, positive and negative count will be assigned to adjective words based on the nature or the polarity of adjective words.

Our system that could summarize the feedbacks, extracting the opinions from all this information, giving an overall view of the product, that could save time and ease the decision process of the customer.

1.6.1 Objectives

The objectives of proposed system are:

• Scrapping product reviews of products from internet. Analyse and categorize review data.

• Analyse sentiment on dataset from document level (review level).

• Categorization or classification of opinion sentiment into- · Positive · Negative.

• Assign the opinions to respective feature of the product.

• To give a statistical representation of the product reviews.

• The proposed system overcomes various drawbacks of the existing manual process of giving the reviews of products.

Summary

This chapter gives an overview of project, its scope, existing system and its disadvantages, proposed system and objectives of proposed system.

CHAPTER 2

LITERATURE STUDY

2.1 Overview

Literature study is carried out to get all the related information of current project, which is used to get an idea for the enhancement as well as changes that can be made to improve existing approaches.

2.2 Related Papers

[1] The consumer reviews for various products are playing a very important role not only for consumers but also for the firms. A large collection of consumer reviews is now available on the internet. These reviews are very helpful to get quality information about the products. The consumer reviews are used as a feedback by the firms in their product development strategies and consumer relationship management. The consumer reviews contain valuable information still we face difficulties in information navigation due to their disorganized nature. The existing product aspect ranking framework automatically identifies important aspects of products from consumer reviews. There are two important observations to identify important aspects. The large number of consumers usually comments important aspects of the product and the consumers' opinion on those aspects have a great influence on their overall opinion about the product. It uses shallow dependency parser for identifying product aspects and sentiment classifier for determining opinion on those aspects.

Finally, it uses a probability aspect ranking algorithm to infer the importance of aspects and ranks it as per their importance score. In this paper, the experimental results confirm the proposed modified system makes the use of aspect rating to improve the performance of important aspect identification and ranking.

The framework of existing product aspect ranking consists of three components: 1) product aspect identification in consumer review; 2) product aspect sentiment classification; 3) product aspect ranking. The consumer reviews contain much more valuable information about the product. Hence, it firstly identifies the product aspects in a review using shallow dependency parser and then analyses the opinions of the consumer on these product aspects using sentiment classifier. Finally,

it uses the product aspect ranking algorithm to know 198 the importance of aspects. For that, it exploits the aspect frequency and influence of opinions which are given by consumers to every aspect over the overall opinions.

[2] With the rapid development of e-commerce, a wide variety of product reviews have appeared on the internet. These reviews not only provide consumers with a reference, but also help manufacturers make reasonable marketing decisions. In online reviews, customers usually give opinions on multiple attributes of products, therefore, the analysis of

product attributes are a crucial issue in product review analysis. This paper studies in depth the extraction and classification of product attribute words from the context. Aiming at the colloquial speech in the online review and the incompleteness of the existing dictionary-based word segmentation methods, this paper uses machine learning method to identify product attribute words.

By introducing the word internal tag method to identify the segmented out-of-vocabulary attribute words, and add it to the user's dictionary, correcting the word segmentation results. In addition, a word-level text classification method based on distributed word representation is proposed, and the semantic and syntactic features of the word vectors are used to classify the product attribute words.

The online product reviews are usually short texts, and a large number of studies have shown that the Condition Random Fields (CRFs) model has good performance in the short text research field, so this paper uses the CRFs model to extract the product attributes. Due to the colloquialization of the online review text and the incompleteness of the existing dictionaries, some of the attribute words cannot be identified. These words may be cut into multiple units after segmentation, such as Chinese words “trunk”, after segmentation, the word is divided into two units “back” and “box” base on Chinese meaning, but in the actual expression of these two units should be put together to express the attribute. Below will introduce the word internal tag method to identify the segmented attribute words. The task of product attribute extraction can be viewed as a process of finding the most likely sequence of tags from the given observation sequence.

[3] Online shopping is more and more common nowadays. The growth in its popularity has led to increase in customer reviews that a product receives. A customer

who has to choose the right product among the huge varieties of products, depends heavily on the product reviews to make a purchase decision. With great volume of product reviews, it become difficult for customers to wade through all reviews to make an informed product choice. Nowadays customers look for features that can serve them specifically. But from the thousands of reviews, it is practically impossible for customers to identify the reviews which speak about the specific product feature.

As a solution to these problems, in this work we aim to analyse a product at feature level, from the customer product reviews. The proposed system, follows a semantic based approach to extract product features. An algorithm, which employ typed dependencies, is introduced for this purpose. Recursive Deep model is used to identify sentiment orientation of review sentences. A review matrix is constructed to find the importance and polarity of each product feature. The experimental results show that the method proposed is effective and has achieved the desired objective.

[4] Review summarization is a process of extracting and collecting reviews which posted on sites. Review summarization helps to gain important information about any product on a less time. This system is mainly suitable for customer and marketer. It provides a platform to see all reviews and buy the product, it is very easy to maintain all the records of a reviews. From consumer reviews important aspects are identified and, on that aspect, provide sentiment classification, and finally apply the ranking algorithm to determine the particular product ranking.

The reviews are very important criteria while purchasing product, because customer take a decision by considering important aspects and firm will concentrate on important aspect while improving the quality. The reviews on product will be positive review or negative review or it will be neutral.

Let's see a sample review "The battery life of blackberry curve is amazing", "The camera is very good of blackberry", "The battery life of Samsung is not good". The above reviews contain opinion about battery life and camera clarity. The first review about battery is a positive opinion and the last review about battery is a negative opinion. About the camera given opinion is positive opinion. There are number of reviews on millions of products that are posted on various websites. CNet.com, Pricegrabes.com is the famous web sites that contain millions of reviews on various products. The consumer can post their review on these websites. And these reviews are useful for consumers who are going to buy new product. Generally,

product may have numbers of aspect. For example, mobile has aspect like design, applications, network, battery etc. for laptop, the aspect such as hard disk, RAM, Graphics card, screen, Battery etc. The product aspects are greatly affecting on product quality

Summary

This chapter discuss about the literature study for the proposed system, related paper which gives idea of enhancements to the existing work. The related paper gives information about the existing systems, its advantages and disadvantages.

CHAPTER 3

SYSTEM REQUIREMENT

The system requirements specify features, components and behaviour of system which is to be developed. The following sections describe about functional, non-functional, performance related, features and behaviour of the solution. This includes the detailed description of the solution to be developed.

3.1 Functional Requirement

Functional Requirements are those requirements which show the working and functionality of a system and the expected behaviour of a system based on certain situations and inputs. It defines specific functionality of a system. Functional requirements of system are:

- Collecting data set
- Normalizing and standardizing the data set.
- Predicting the value from the data set.

3.2 Non-functional Requirement

Non-functional requirements is not about functionality or behaviour of system, but rather are used to specify the capacity of a system. They are more related to properties of system such as quality, reliability and quick response time. Non- functional requirements come up via customer needs, because of budget, interoperability need such as software and hardware requirement, organizational policies or due to some external factors.

3.2.1 Basic Operational Requirement

The primary functions of systems engineering are all performed by the end users, which is the customers. Operational requirements which are given by:-

- **Reliability**: The reliability of the product will be dependent on the accuracy of the user comments, the product searched and the features of the particular product.

- **Security:** The user will only be able to access the website using his login details and will not be able to access the computations happening at the back end.

- **Maintainability:** The maintenance of the product would require training of the software by recent data so that their commendations are up to date. The database has to be updated with the user comments.

- **Portability:** The complete project will be portable as it is written in java which is portable. The java code is written once and it runs on any JVM.

- **Ease of Use:** The user interface allows the user to interact with the system at a very comfortable level with no hassles.

- **Modularity:** The many different modules in the system are neatly defined for ease of use and to make the product as flexible as possible with different permutations and combinations.

- **Robustness:** During the development of the system special care is being taken to make sure that the end results are optimized to the highest level and the results are relevant and validated. Java language is used for the development, itself provides robustness to the system and thus makes it highly unlikely to fail.

3.2.2 Organizational Requirement

The Organizational requirement consists of the following types:

- **Process Standards:** To make sure the system is a quality product, IEEE standards have been used during system development.

- **Design Methods:** Design is an important step, on which all other steps in the engineering process are based on. It takes the project from a theoretical idea to an actual product. It gives us the basis of our solution. Because all the steps after designing are based on the design itself, this step affects the quality of the product and is a major player in how the testing and maintenance of a project take place and how successful they are.

3.3 Resource Requirement

1. **Eclipse:** It is an integrated development environment (IDE) used in computer programming, and is the most widely used Java IDE. It contains a base workspace and an extensible plug-in system for customizing the environment. Eclipse is written mostly in Java and its primary use is for developing Java applications. Eclipse uses plug-ins to provide all the functionality within and on top of the run-time system. Its run-time system is based on Equinox, an implementation of the OSGi core framework specification. Advantages of Eclipse are: Code Completion, instead of digging through documentation you should be able to tab your way through methods and save yourself a lot of writing.

2. **SQLyog :** It is a GUI tool for the RDBMS MySQL. It is developed by Webyog, Inc. based in Bangalore, India and Santa Clara California. SQLyog is being used by more than 30,000 customers worldwide.

   Prominent features of SQLyog are:

   - 64 bit binaries are available from version 11.0.
   - Editor with syntax highlighting and various automatic formatting options.
   - Intelligent Code Completion.
   - Data manipulations (INSERT, UPDATE, DELETE) may be done from a spreadsheet-like interface. Both raw table data and a result set from a query can be manipulated.
   - Visual Schema Designer.
   - Visual Query Builder.
   - Query Formatter.

3.4 Hardware Requirement

The following is the hardware requirements of the system for the proposed system:

- Processor : 2.2 GHz
- RAM : 4GB
- Hard Disk : 50 GB

3.5 Software Requirement

The following is the software requirements of the system for the proposed system:

- OS : Windows 8 and more updated version.
- Platform : Any Browser.
- Language : Java
- IDE/tool : Eclipse

Summary

The chapter describes about functional, non-functional, resource, hardware and software requirement of system.

CHAPTER 4

SYSTEM ANALYSIS

Analysis is nothing but finding solution to various problems. System analysis is defined as, process in which we get the information about the existing problems, requirements and to solve various problems related to system. Study of feasibility plays vital role in system analysis, which helps in providing goals for development and design.

4.1 Feasibility Analysis

Considering initial results, a deeper survey or feasibility study is done on existing results. “FEASIBILITY ANALYSIS” is nothing but meeting the requirement of system, to achieve desired goal, working on proposed system and getting detailed information on resources.

There are 8 steps to be considered for Feasibility Analysis:

- Develop a team for a project and assign leader for that team.
- Identify the strength of proposed work.
- Characterize and recognize qualities of proposed work.
- Identify cost of proposed work and show its performance.
- Show system performance.
- Choose best framework.
- Prepare report and submit to administrative

Three key features involve in “FEASIBILITY ANALYSIS” are:

- Financial Feasibility
- Technical Feasibility
- Public Feasibility

4.1.1 FINANCIAL

Financial feasibility is one of key factor of feasibility analysis, which is carried out to scrutinize economic cost of organization, limited fund is endowed by companies for development of system. Hence, the developed project cost is within budget, because of freely available resources. Only personalised components need to be bought.

4.1.2 TECHNICAL

Technical feasibility is entailed to scrutinize technical performance of system. System which is building should have less demand. Hence, this gives number high of demands asked by the clients. The proposed system is built in such way that there is no harm for the user .

4.1.3 PUBLIC

The proposed system should be developed in such a way that user should accept all the necessity and should not get threaten by the system. The proposed system is developed in such an efficient manner, which gives more efficient, value of open and close values which can be used in many real world datasets.

Summary

The objective of this chapter is to know the proposed system is feasible or not. The chapter mainly describe about the various keys of feasibility analysis i.e. financial, technical and public.

CHAPTER 5

SYSTEM DESIGN

System design is a process of giving detailed information about the proposed work in a physical format. Different designs are built for development of system, which describes about features, components which are included and how client interact with system.

5.1 Fundamental Design Concepts

Fundamental design is developed in course of recent years. As year passes, enthusiasm of creating new designs is evolved and each design has been tested. Software designer gets new ideas and foundation to build and test new design concepts. Fundamental framework is design to "getting it right". Major plan ideas, for example, deliberation, refinement, modularity, programming engineering and data encryption is applied to meet the requirement of proposed work

5.1.1 Input Design

Input Design is a way toward changing user-based inputs into computerized format. Main objective of input design is, to make computerization as possible and error free. Giving a decent information configuration to the application simple information and determination highlights are received. The input design prerequisites, for example, ease of use, reliable organization and intelligent exchange to help client to get proper information on time. Input design is a general framework which exceptionally cautious consideration. Gathering all input parts is one of the costly parts of framework.

5.1.2 Output Design

Output design meets the necessities of client and presents the output data clearly. In any framework processing result are conveyed to clients and different frameworks in form of output design. It is direct source to client. Productive output enhances framework association with machines of source and destination. Output is an optimized statistical display of all summarised reviews.

5.2 Development of System

Development of system is a method, where development of product is completed or it solves all problems of system. Development of software includes number of stages and process to develop software. Step by step procedure is followed to complete development of software. The method which is followed in this project is waterfall model.

5.2.1 Different phases of model

Requirement: This phase includes collection of all requirements which is needed for development of software.

Design: The specification of system is converted into software design, by keeping in mind system specification. The designer mainly describes about algorithm, architecture and structure of system.

Coding: Developer begins coding with a specific end goal to give a full outline of project. As such framework system specification are just changed over into system decipherable process code.

Implementation: This stage includes execution of project which involves coding. The output is commonly a library, documentation and client manuals.

Testing: The testing stage includes all modules of project which is coordinated and tried to guarantee that total framework meets product necessities. Verification and Validation is mostly concerned under this testing phase.

Maintenance: It is most important stage, where product should be user friendly, adaptable, error free and should improve productivity of project.

5.2.2 Purpose of choosing Water fall model

- Objectives of project should be clear.
- Understandable software requirement.
- System progress should be measurable.
- Allocation of resources should be better.

- Guides to be prefect.
- Stable project requirements.
- Quality of project is improved.

5.3 System Architecture

The System architecture for the proposed system is shown below:

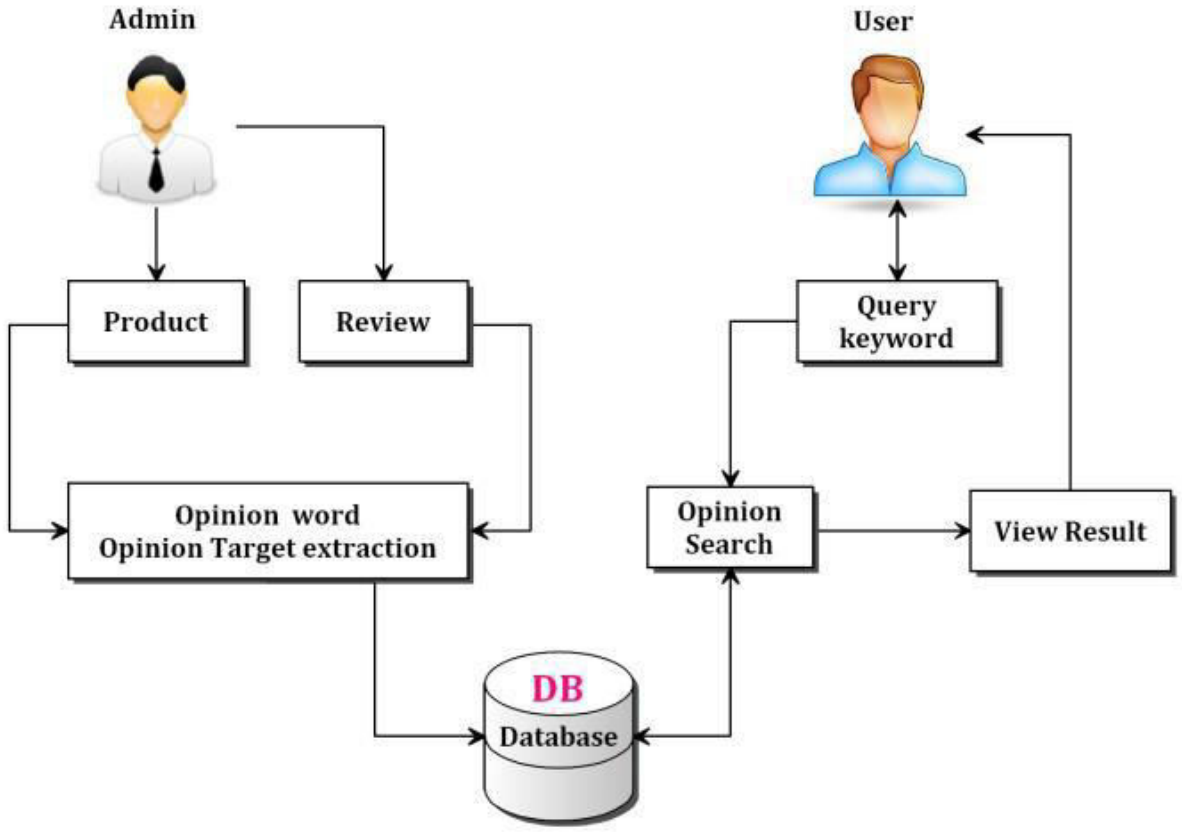

Fig 5.1: Proposed System Architecture

The modules in this project are shown below:

- **Adding reviews from dataset:**
  In this module we are going to feed offline reviews based on the category id of the product. Based on the product id admin going to feed the offline reviews and going to maintain in the database.

m_review

Review_no	Pro_code	Review_comment
101	Iphone 5	This phone has a colorful and big screen
102	Iphone 5	This phone has a colorful and big screen but its LCD resolution is very disappointing.
103	I10	Pick up is good for this car
104	I20	Look of the car is excellent but car mileage is not good

Fig 5.2: Dataset

Admin is selecting iPhone 5 in dataset and extracting the following records

101	Iphone 5	This phone has a colorful and big screen
102	Iphone 5	This phone has a colorful and big screen but its LCD resolution is very disappointing.

Fig 5.3: Extracted reviews

- **Removal of unnecessary words:** For Each comment remove unnecessary words.

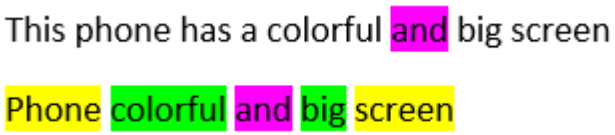

Fig 5.4: Removing unnecessary words

Here the word ‘and’ is unnecessary so it will be removed.

- **Identifying noun, adjective and classifying positive and negative words using NLP:**

  In order to gain the sentiment of the user, these sentences have to be segregated into words in which the adjectives, nouns and adverbs are processed using Natural Language Processing (NLP). Nouns will be considered as the opinion target or the feature and adjectives will be considered as opinion words.

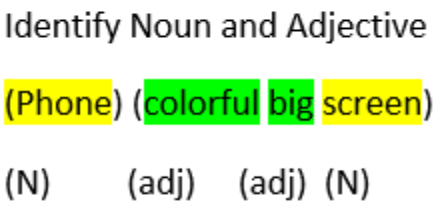

Fig 5.5: Identifying noun, adjective

Finally, positive and negative count will be assigned to adjective words based on the nature or the polarity of adjective words.

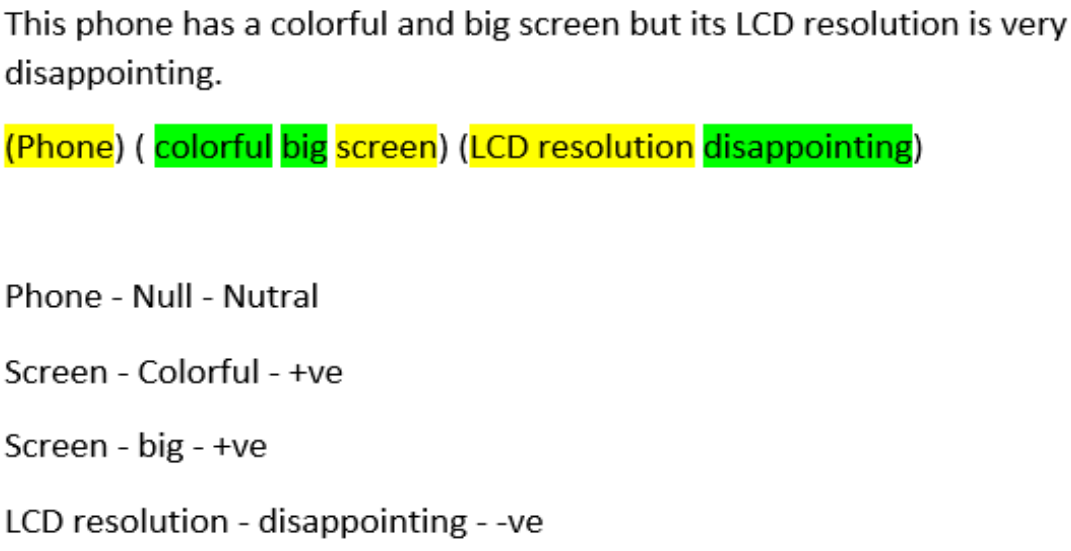

Fig 5.6: classifying positive and negative words

- **Mapping Opinion Target-Opinion Word pair and assigning weightage:**

  In this module we are going to extract opinion word and opinion target. Based on the extraction of words we will make noun and adjective pair. So, opinion target will be mapped with the respective opinion word.

Op_pair_no	Opinion_target	Opinion_word	Pro_code	Op_pair_value
1	Phone	Null	Iphone 5	0
2	Screen	Colorful	Iphone 5	1
3	Screen	Big	Iphone 5	1
4	Phone	Null	Iphone 5	0
5	Screen	Colorful	Iphone 5	1
6	Screen	Big	Iphonr 5	1
7	LCD resolution	Disappointing	Iphone 5	-1

Phone - Null - 2 - 0
Screen- big - 2 - 2
Screen-colorful - 2 - 2
LCD Resolution - disappointing - 1 -(-1)

Total Value : 4 -1 = 3

Fig 5.7: Mapping Opinion Target-Opinion Word

Finally, we will assign positive and negative count based on the adjective words.

Use case Diagram

Use case diagram is a representation of a user interaction with the system that shows the relationship between the user and the different use cases in which user is involved.

Admin maintains following modules:

- Login Module
- Category Module (View Only)
- Product Module (Add, Edit, Delete, View)

- Review Module (View Only)
- OT-OW Extraction
- View OT-OW (Product Wise)
- Change Password

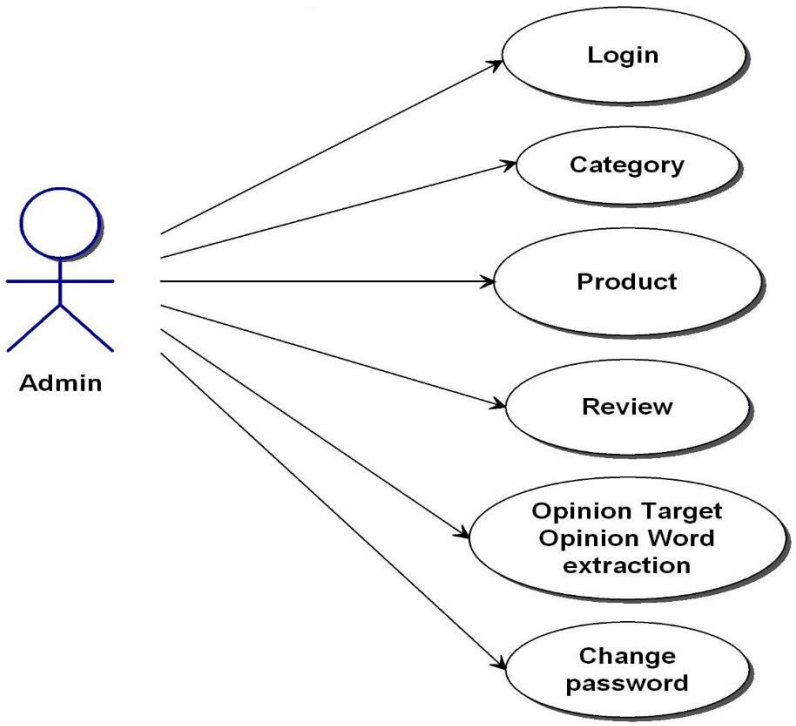

Fig 5.8: Use Case Diagram- Admin Session

Users deal with the following modules:

- Member Registration.
- Login Module.
- OT-OW extraction module.
- View all the polarity pair with weights.
- Change Password.

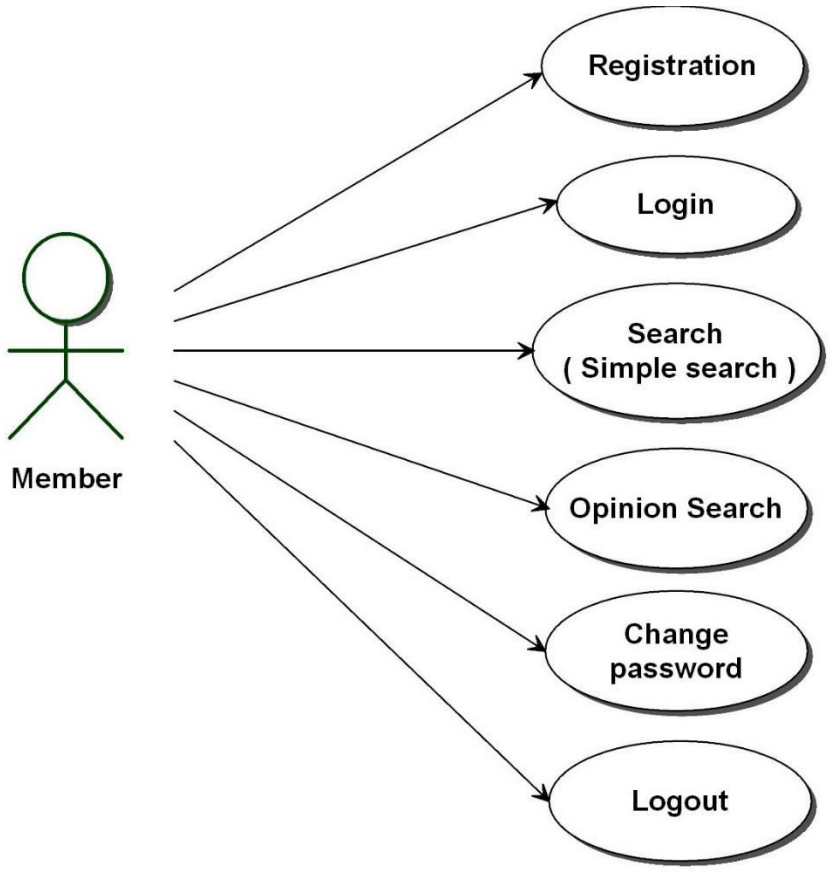

Fig 5.9: Use Case Diagram- Member Session

5.4 Sequence Diagram

Sequence diagram is sort of interaction diagram which describe about process of work done to execute project.

The following shows the sequence diagram for initialization:

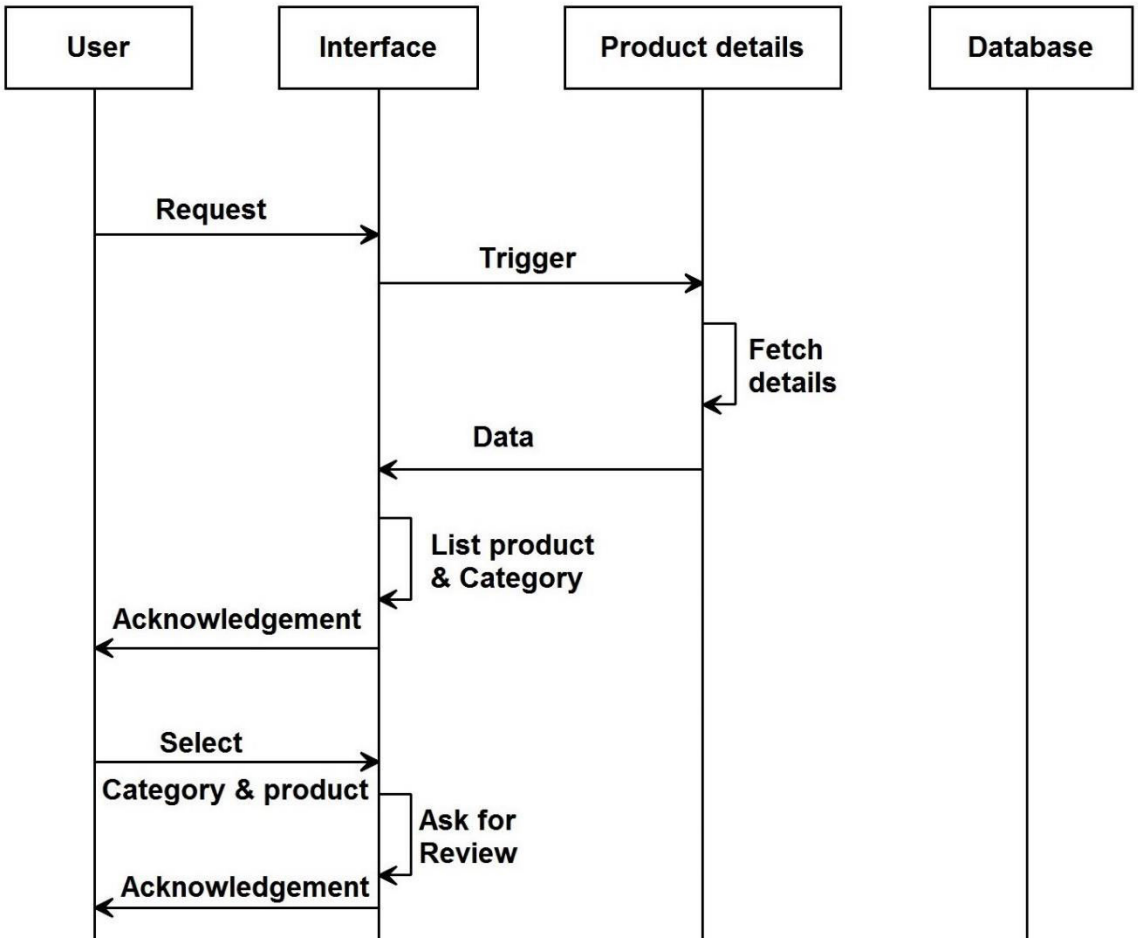

Fig 5.10: Sequence diagram-user

- User requests product details.
- Interface fetches the required details and returns the data.
- List of categorised products will be displayed.
- User chooses his/her required product.

The following shows the sequence diagram after user chooses his/her required product.

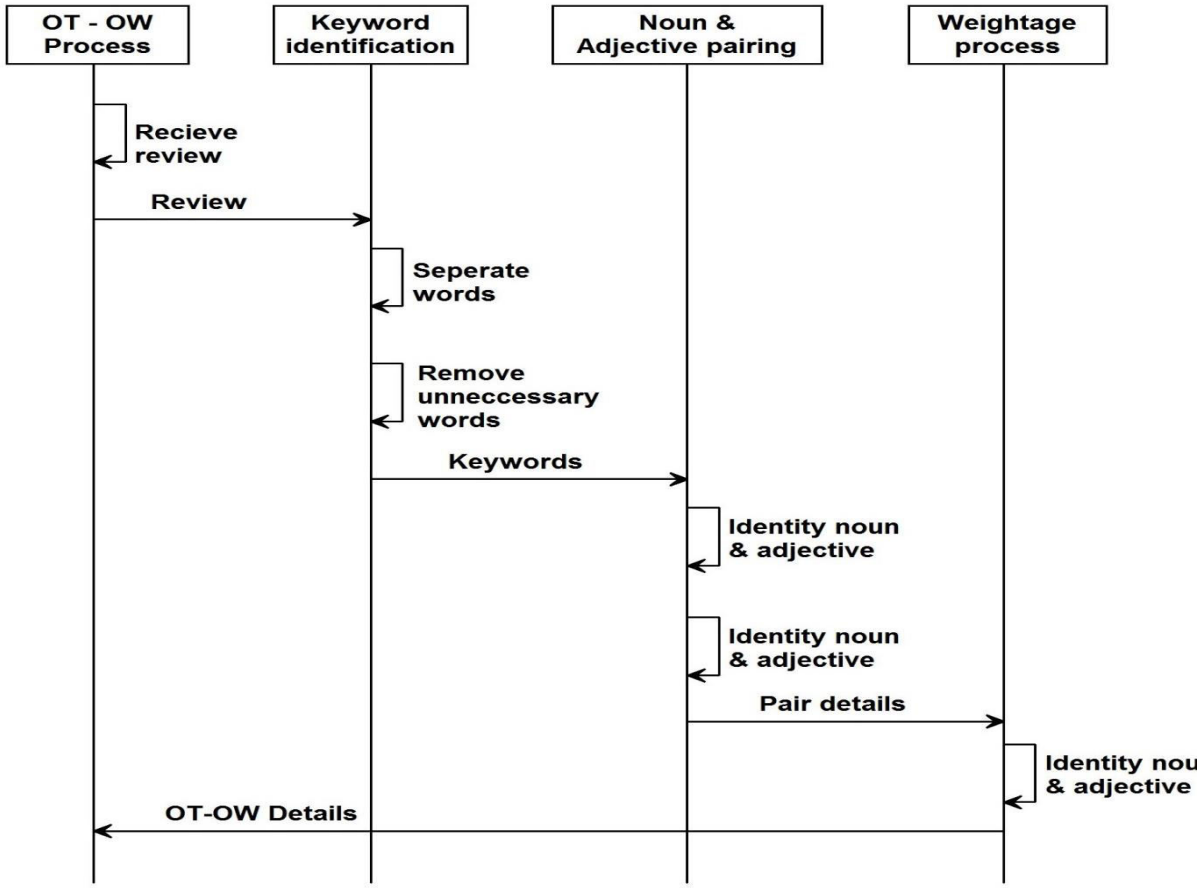

Fig 5.11: Sequence diagram for OT-OW process

- Using the reviews stored in the database, opinion target and opinion word are mapped.
- Keyword identification is done and unnecessary words are removed.
- From identified keywords noun and adjective pairing is done.
- Based on the paired details weightage is assigned.
- OT-OW details are updated.

Summary

This chapter mainly concentrates on input design, output design, waterfall model, system architecture and various designs which is involved in proposed system.

CHAPTER 6

IMPLEMENTATION

It is main part of project where project is implemented. Implementation must be clearly defined, planned carefully and systematically otherwise it causes confusion and leads to generation of problems. Following tasks involved in stage of implementation:

- Planning carefully
- System Investigation.
- Investigating different types of systems and constraints.
- Selecting the most accurate and correct language for developing the application.
- Understanding and correctly evaluating several changeover methods.
- Rightful decision is made about the selection of the model.

6.1 Language used for implementation

Implementation is an important step where execution of project is done. Hence language used for implementation should be adaptable, error free, user friendly and etc. Most of the time because of some mistakes project gets ruined or spoiled due to inappropriate programming language.

Here JSP (Java Server Pages) is selected as the front end which is based on HTML and XML and CSS is used for styling. For the back end Java is used as the preferred programming language for implementation of project along with MySql programming language used for storage in the database. Reasons for selecting Java are given by:

1. Platform Independent: Java is supported by many platforms like Linux, Windows, and Macintosh etc. Because the byte code can we run on any JVM. Hence, Java is platform independent language used in various applications.

2. Object Oriented: Java is based on objects which is class based. Hence Java is object-oriented language i.e. everything in java is an object.

3. Standard Library: The Java Class Library is the standard library, developed to support application development in Java. This is in cooperation with Java Community Process

program .Companies or individuals participating in this process can influence the design and development of API's.

4. Security: The Java platform provides a number of features designed for improving the **security** of Java applications. This includes enforcing runtime constraints through the use of the Java Virtual Machine (JVM), a security manager that sandboxes untrusted code from the rest of the operating system, and a suite of security APIs that Java developers can utilise.

5. Garbage collection: garbage collection is a form of automatic memory management. This attempts to reclaim *garbage*, or memory occupied by objects that are no longer in use by the program. Hence this saves space and in turn makes program easy to execute without any programming errors and memory errors.

6.2 Implementation platform

Platform plays an important role for development of software. A platform is a place where software is launched. Windows platform is used for implementation of project. There are various analyses for choosing Windows platform. It is used for many different purposes such as scrutiny, embedded system, security etc. it also contains remote connection and option to restore, so that file can be restored if deleted. Windows software is used for scalable processing, multi-tasking, Encrypted File System (EFS) and smart card support.

Summary

The chapter describes about implementation process of proposed system. It describes about which platform, language and processing techniques used for executing the project.

CHAPTER 7

TESTING

Testing is an important part of project where testing of each module is done. Testing guarantees that proposed system is well organized analysed to meet require project goal. Testing is last stage of project which guarantees the system is error free and ready to give desired output. The goals of testing are given by:

- Give operational quality to system
- Search and remove errors.
- Best quality project is produced.
- To approve the product as a solution for the first issue.

The following are types of testing performed in proposed system.

7.1 Unit Testing

Each module of project is tested individually. Verification is done on each module. Module of project is tested individually. Testing is done in programming style. The unit testing for the proposed system is performed on initialization, data importing, data standardization and prediction.

7.2 Integration

After unit testing is performed, integration testing takes place in each module of project. Integration is done on various classes of system. Integration testing is done on front and back end also.

➢ Function into classes Integration

Initially during code phase various functions is developed for development of system. Each function of system is tested and coded individually. As all the functions are verified they are mixed into their particular classes.

- **Distinct classes Integration**

  Based on functionality, testing of distinct classes is done independently. Verification of each class is done which gives good result and hence integration is performed again on different classes.

- **Front & Back end Integration**

  Java along with JSP is used for front end. MySQL is used for backend, using database connection frontend and backend are connected. Hence backend is integrated with GUI and later tested.

7.3 Integration Testing

Developing a programming framework is a sophisticated technique which is used by Integration testing. It solves various issues on dual verification problem and construct program which solves all related problems. Main objective of integration testing is to construct a program structure based on unit testing modules.

As modules of software are divided, testing is performed on each module. Later this separated module is tested as whole set. Here, to rectify errors is difficult as it has different isolated errors.

7.3.1 Up down Integration

The up down integration deals with development of program framework in incremental way. Modules of each program are co-ordinated in descending order and it starts with primary module.

7.3.2. Middle-up Integration

The below table integrated testing table is divided into integrated classes, functions of each classes, how test is performed and result generated. It is important to check whether the testing is error free or not for different classes.

Integrated class and functions	Input	X – Expected Result Y - Actual Result	Testing	Observation
Class : Textsearch text_filter() text_Search()	Battery is Good.	X = [Battery , Good] Y = [Battery , Good]	To check class test whether all commands are working properly or not.	Success
Class : OTOWExtract getReview() getOTOW()	Battery is Good.	X => Battery = 1 Y => Battery = 1	To check class test whether all commands are working properly or not.	Success
Class: AdminDAO addReviewDetails() getProfile()	Select * from reviews	X => Reviews Y => Reviews	To check class test whether all commands are working properly or not.	Success
Class: UserLogin doPost()	UserName Password	X = Login successful Y = Login Successful	To check class test whether all commands are working properly or not.	Success

Table 7.1: Integration testing table

7.4 Validation Testing

As completion of testing combination is done, writing computer program is put together in on package. Testing approval is described from various perspectives. Hence testing affirms items limits which are sensibly expected.

Functionality which is tested	Inputs	Testing	Observation
Working of Adding dataset.	The input fed is the user comments, that is extracted from amazon.	Testing whether the added file is uploaded successfully and data is present in database.	Success
Working of OTOW extraction.	The extracted nouns and adjectives are mapped to the words present in the database.	Nouns andadjectives were extracted from user comments and mapping is done.	Success
Working of final result to the user	User has to click on particular product to get the scores of each feature of the product.	The scores for each feature of the product will be displayed for the user	Success

Table 7.1: Validation testing table

7.5 Testing Output:

Output testing of proposed system is performed after the validation testing where the system produces required output in specified format. The output will be displayed both in the GUI and the eclipse console. The output generated for each and every product is manually tested.

7.6 User Acceptance Testing:

Key factor for success of any system is user acceptance. User acceptance testing is performed on users, show which will be successful based on user motivation and knowledge. At time of developing and making required changes, system under consideration along with prospective system, users undergoes constant testing. The changes are made regarding to 3 points

- Input design
- Output design
- Menu driven system

7.5.1 White Box testing:

White box designs test cases according to internal structure by considering an internal perspective of the system. It identifies all paths through the software. The output is determined based on the chosen test case inputs.

Test cases derived by white box testing ensures

- Within a module all the independent paths are tested once.
- Logical decision testing on right and wrong sides.
- Loops are executed operational bounds.
- Validity is assured by executing the internal structure.

7.5.2 Black box testing

Functional testing is considered as Black box testing and it mainly focus on functional requirements. It is a software technique where the tester could not predict the testing of internal workings. Programming code is not examined by the tester and they need not to have further knowledge of the programming other than its specifications. It is a complementary approach uncovering distinct classes of errors which are:

- Error in performance
- Initialization and termination errors

- Missing function
- Errors in interfaces
- Errors in objects

Advantages of black box testing:

- Tests are unbiased
- Programming knowledge is not required.
- Based on user point of view test is conducted.
- Cases are built after completion of specifications.

7.7 Test Data Preparation

In system testing, test data preparation plays a main role. Test data is prepared and is used to test the system under study where errors are hidden and removes by using following testing steps for better future use.

7.8 Assurance of Quality

Quality assurance is testing and analysis of administration element. The main objective is to give knowledge about item quality to administration. Assurance of quality involves:

7.8.1 Quality Factors

The main goal of confirmation value is track product quality and observes procedures to enhance programming. Quality factors are described in following two categories:

- Directly measured factors.
- Indirectly measured factors.

Quality factors mainly focus on following three things:

- Operational attributes
- Experiences capacity changes
- Versatility.

- Effectiveness
- Time duration.

7.8.2 General Risks

Risk is nothing, but which gives negative results at undesirable incident. The following three things are considered to recognize risk in other projects:

- Damage occurs during an occasion.
- Probability of occurring an occasion.
- The result that modified at certain level.

7.8.3 Security Technologies & Policies

The seven major activities the software quality is comprised of which are follows:

- Conduct of formal specialized audits
- Measurement
- Testing Software
- Application for specialized techniques.
- Control of progress
- Record keeping and announcing
- Enforcement of measures

Summary

The chapter describes about various testing techniques such as unit testing, integration testing, validation testing, output testing and etc.

CHAPTER 8

RESULTS

8.1 Result Screenshot

8.1.1 Admin Side

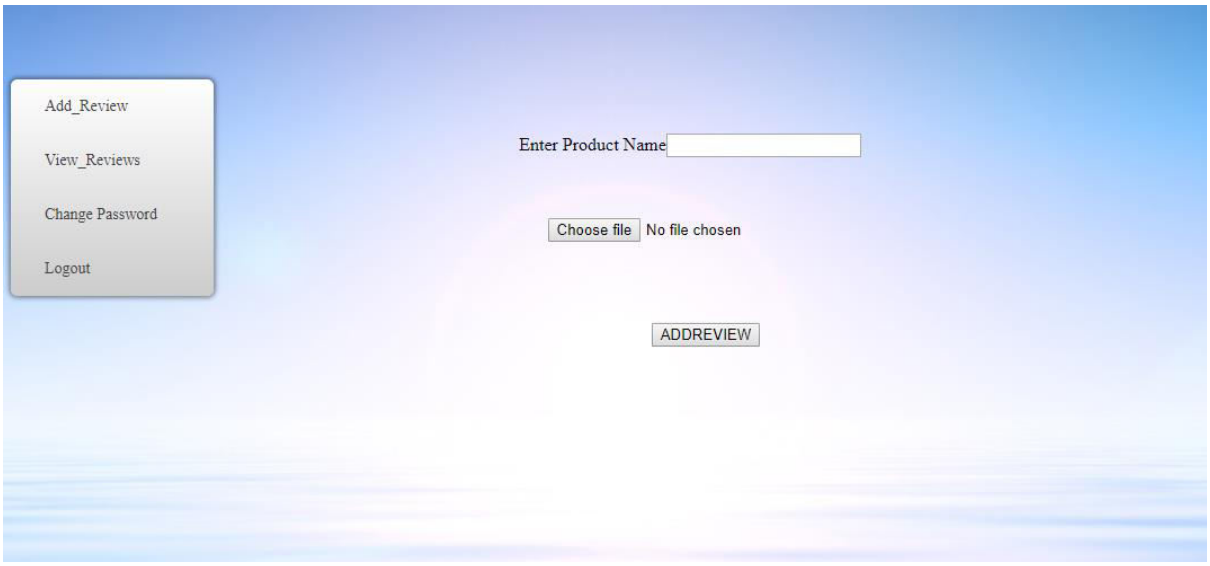

Fig 8.1: Admin adding dataset

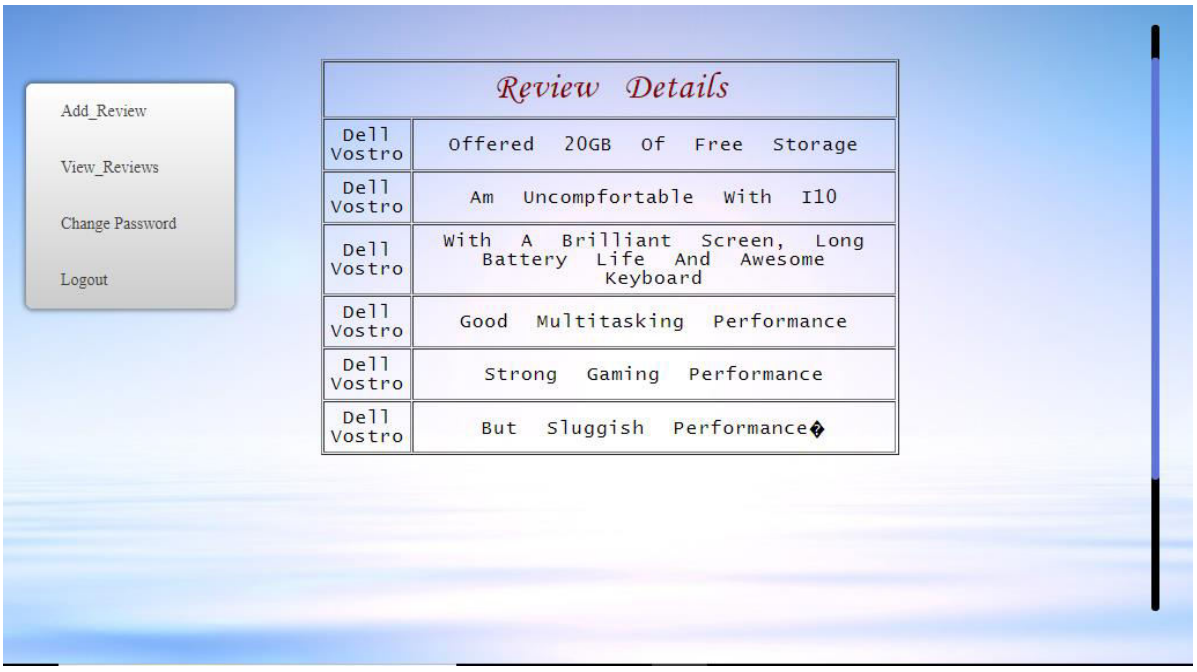

Fig 8.2: admin viewing data

Fig 8.3 Admin changing Password

8.1.2 User Side

Fig 8.4 User Registration

Fig 8.5: User can search reviews of a particular product

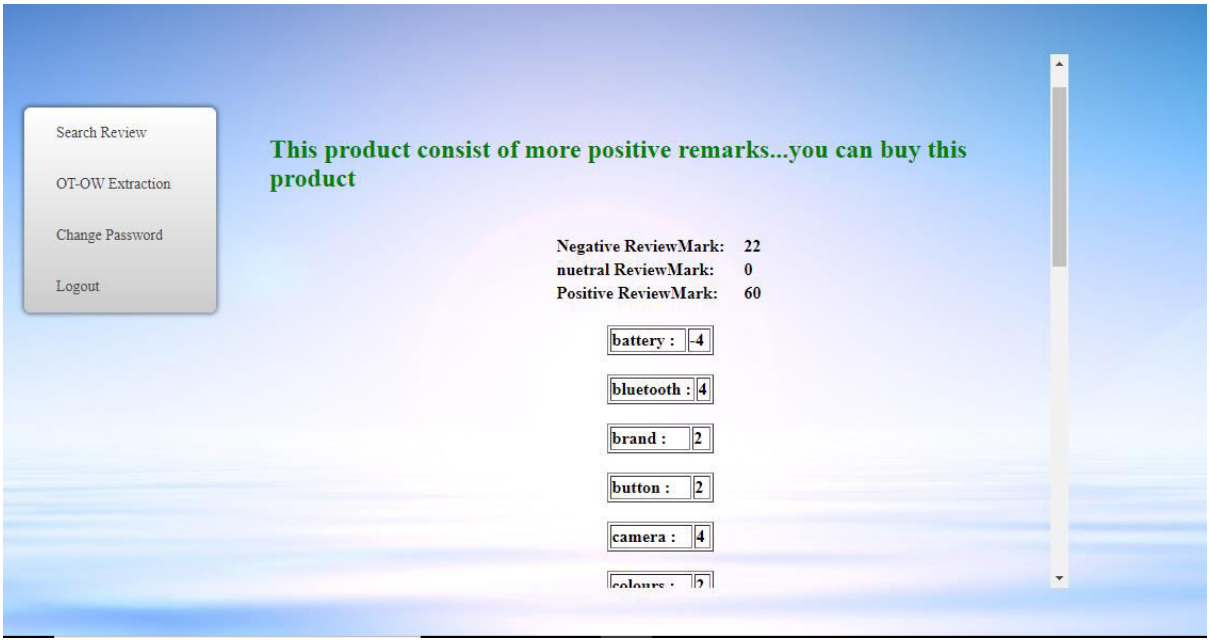

Fig 8.6: Final result

In the Fig. 8.1,8.2 the admin adds the data set for a particular product by entering its name and choosing the file for the product entered and later clicks on the add review. Here a

series of steps take place where xlsx file is read and formatting takes to remove whitespaces and each line is parsed and based on the product name the data is stored review by review in to the database. Hence all the reviews are read and stored on two columns. Later the admin can also view the various data stored in the database which is fetched into the Interface this action takes place when the admin clicks on the view reviews. Here the reviews of all the products entered at that moment or previously added reviews are also fetched from the database into the user interface. In Addition to these main processes a new admin can himself register and gain accesses to this side of the web application.

In the Fig. 8.5, 8.6 the screen shots related to the user side of the web application if the user wishes to view the reviews about any product he can search for the same and view the reviews if they are available, If present the data will be retrieved from the database and displayed in the user Interface. The results are provided when the user searches for the product that he/she requires from the drop down list. When the user selects the available product a series of tasks take place.

The first of the tasks taking place is searching for all the reviews stored in the database for the particular product and then removing various unnecessary words and formatting the reviews to a standard format also removing whitespaces, symbols .Then mapping of ot-ow (opinion target, opinion word)by identifying nouns and adjectives.

Later identification of the positive and negative opinion about the features is recognised (adjectives) and are stored in the database. A positive score is assigned if the opinion is positive and similarly a negative score is added otherwise. This process is carried out for each review and hence an adaptive feature score is provided for the reviews covering the particular feature.

8.2 Major Advantages of Proposed System

1. Scores are provided for each feature of the product.

2. This application adapts to different variety of products.

3. Since all the features are not covered in the review it automatically adapts to the specifications covered.

4. Since scores are generated from reviews they are free from bias.

Summary

The chapter describes about various screenshot of results, step by step process to execute proposed project and advantages of proposed system

CHAPTER 9

CONCLUSION AND FUTURE WORK

9.1 CONCLUSION

Giving scores to every feature of the product gives customers a summarized view and hence saves the customer from bulky product reviews and therefore achieves the goal of feature based analysis for all products. This method has been found to be largely helpful to customers and helps them take decisions effectively.

Analyzing the review given and converting them into scores is possible only for direct sentences and commonly used words but the code fails to analyze sarcasm present in the reviews it also fails to link references from a context happening in the present world. The data set used is available from third party sites which gain data from e-commerce sites hence availability of the data set for a particular product is dependent on such third-party websites.

9.2 FUTURE WORK

Considering the future scope of the proposed system, we would consider gaining accesses from e-commerce websites to dynamically accept the reviews given by the customers in the particular site and also combine data sets from various sites and hence gain an aggregated data from all the sites which would have added advantages such as avoiding bias for a particular brand by the e-commerce sites. Hence the future proposed system would probably also include identifying sarcasm and references to context.

As future work we propose to offer a summary of reviews for more than 2 products and also automatically rank products based on the features that the user is interested in. And we will implement an algorithm to take care of user reviews which consists of sarcasm. We will try to decrease the time consumption for getting the results.

REFERENCES

[1] Recommending Products to Customers using Opinion Mining of Online Product Reviews and Features. 2015 International Conference on Circuit, Power and Computing Technologies [ICCPCT].

[2] Sentiment analysis from product reviews using SentiWordNet as lexical resource. ECAI 2015 - International Conference – 7th Edition Electronics, Computers and Artificial Intelligence 25 June 27 June, 2015, Bucharest, ROMÂNIA.

[3] Feature Extracted Sentiment Analysis of Customer Product Reviews. 2016 International Conference on Emerging Technological Trends [ICETT].

[4] Aspect Rating Analysis Based Product Ranking. 2016 International Conference on Global Trends in Signal Processing, Information Computing and Communication.

[5] Online Product Review Summarization. 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS). [6] Datasets : https://www.kaggle.com/PromptCloudHQ/amazo n-reviews-unlocked-mobile-phones.

[5] OSHA (2012). https://osha.europa.eu/en/about-eu-osha/national-focal-points/turkey (Last Access date: 15.04.14).

[6] Pang, Bo, and Lillian Lee. "Opinion Mining And Sentiment Analysis." Foundations and trends in information retrieval 2.1-2 (2008): 1-135.

[7] Wang, Hao, et al. "A System For Real-Time Twitter Sentiment Analysis Of 2012 Us Presidential Election Cycle." Proceedings of the ACL 2012 System Demonstrations. Association for Computational Linguistics, 2012.J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[8] Balahur, Alexandra, et al. "Opinion Mining On Newspaper Quotations." Web Intelligence and Intelligent Agent Technologies, 2009. WI-IAT'09. IEEE/WIC/ACM International Joint Conferences on. Vol. 3. IET, 2009.

[9] Pak, Alexander, and Patrick Paroubek. "Twitter As A Corpus For Sentiment Analysis And Opinion Mining." LREC. 2010.

[10] Bollen, Johan, Huina Mao, and Xiaojun Zeng. "Twitter Mood Predicts The Stock Market." Journal of Computational Science 2.1 (2011): 1-8.